DOCUMENT RESUME

ED 197 561 EC 131 768

AUTHOR Gersten, Russell M.; And Others
TITLE Reconceptualizing Vocational Assessment for the Severely Retarded: Evolution of the Vocational Performance Sample.
PUB DATE Apr 80
NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April, 1980). Best copy available.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adults; Criterion Referenced Tests; *Occupational Tests; *Severe Mental Retardation; Test Reliability; Test Validity; *Vocational Aptitude
IDENTIFIERS *Vocational Performance Sample

ABSTRACT

The Vocational Performance Sample (VPS), a vocational assessment instrument for the severely retarded, integrates norm referenced technology with current innovations in domain referenced tests and learning potential assessment. The domains are based on a systematic sampling of types of training procedures and types of discriminations and/or operations currently used in vocational training for the severely handicapped. Correction procedures are integrated into the test format. The test was validated on 158 severely and profoundly retarded adults. Internal consistency was .92; test retest reliability was .92. VPS scores correlated moderately well with both a Supervisor's Checklist Scale and raw scores on the Peabody Picture Vocabulary Test. Criterion related validity was high with VPS scores accounting for over 50% of the variance in training time on a vocational task. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED197561

BEST COPY AVAILABLE

Reconceptualizing Vocational Assessment for the Severely Retarded: Evolution of the Vocational Performance Sample

Russell M. Gersten, Thomas J. Heiry, Larry Irvin
University of Oregon

Paper presented at AERA,
Boston, Massachusetts, April, 1980

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Russell Gersten

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EC 131768

Abstract

The Vocational Performance Test, a vocational assessment instrument for the severely retarded, integrates norm-referenced technology with current innovations in domain-referenced tests and learning potential assessment. The domains are based on systematic sampling of types of training procedures and types of discriminations and/or operations currently used in vocational training for the severely handicapped. Correction procedures are integrated into the test format. The test was validated on 158 severely and profoundly retarded adults in Oregon. Internal consistency was .92; test-retest reliability was .92. VPS scores correlated moderately well with both a Supervisor's Checklist Scale and raw scores on the Peabody Picture Vocabulary Test. Criterion-related validity was high with VPS scores accounting for over 50% of the variance in training time on a vocational task.

The purpose of this paper is to describe the evolution of an instrument designed to aid in placement decisions and predict relative training time for severely and profoundly retarded adults. A second, perhaps more important, objective is to delineate how tests can be developed which combine the rigor of psychometric standards usually associated with norm-referenced tests with some of the innovative thinking associated with the domain-referenced test movement and the learning potential assessment models of Budoff and Hamilton (1976) and Feurstein (1972).

Vocational training for the severely retarded was virtually non-existent seven years ago. The use of the principles of applied behavior analysis to this enterprise during the last decade by Gold (1973), Bellamy, Horner & Inman (1979) and others has led to a highly sophisticated technology where individuals with IQs of 12 or so can be taught complex electronics assembly tasks. Yet, little has been done in the area of assessment or measurement of individual differences. Like many other behaviorists, leading authorities in the field such as Gold (1974) and Bellamy et al. (1979) generally argue for idiographic assessment of each client. But there are many reasons why this is infeasible, especially in understaffed, large state institutions where, due to legislation, thousands of placement decisions will be made in the next decade. Our working assumption, like Gold's, is that virtually all individuals can be trained, but that a) training time, and b) appropriate training strategies will vary according to measurable individual differences.

The Vocational Performance Sample was validated and field-tested on a sample of 158 profoundly and severely retarded adults. No normal "control" sample was included. Thus all items are relevant to a severely handicapped population, and training and teaching methodologies for this population.

All the traditional procedures for norm-referenced test development--reliability (normal consistency and temporal stability), concurrent validity and criterion related validity--have been followed, which seems quite appropriate for a test whose primary aims are prediction of training time and assistance in placement decisions. A secondary objective in developing this test was our belief that a carefully validated instrument could offer precise information on vocationally relevant patterns of individual differences and potential information on aptitude-training interactions. In addition, a picture of the strengths and weaknesses of current training programs for the severely retarded may emerge.

Methodology and Results

Content validity was a very serious concern in the construction of the test. An item-generating grid as represented in Illustration 1 was developed using two variables: a) type of training procedure (i.e. model, verbal directions, physical prime) and b) type of discrimination or operation (e.g. coordinated 2-hand movement) required of the subject. This assured that the major components of both Gold's and Bellamy's training programs were adequately covered. Thirty-three items were developed by this procedure.

Insert Illustration 1

The most unique feature of the VPS was the incorporation of a modified version of the learning potential model (Budoff, & Hamilton, 1976; Feurstein, 1972) into the test procedure. Each item was presented to the client twice; if s/he failed the first time a correction procedure (modeled after those used in the forementioned training programs) was introduced. If the subject responded to the correction appropriately, s/he received 1 point, if s/he performed the task correctly twice without need of correction, 2 points were given. Thus the subject's ability to profit from correction was incorporated into the test format. This is particularly important for individuals in institutions who have often been deprived of access to even the most basic education and training.

Internal consistency, measured by coefficient alpha, was .92 (N = 158). Test-retest reliability over a 2 to 3 week interval for 26 randomly selected subjects was .93.

Concurrent validity was assessed using two measures, a newly developed Supervisor's Checklist (Gersten, Irvin, Bellamy, Close, Heiry, 1979) and raw scores on the Peabody Picture Vocabulary Test. The VPS correlated moderately well with the Supervisor Checklist, with a median correlation across 11 work activity sites of .76. Using the total samples (N = 158) the validity coefficient between the VPS and raw scores on the PPVT was .75. The median correlation across sites between the Supervisor Checklist and the PPVT was .40. This pattern of relationships demonstrates that the VPS measures something in common with both of these instruments, but the PPVT and the checklist generally measure things unique to each.

Another outcome was that the correlations between supervisor's ratings and the VPS were generally much higher at small community sites than at

larger institutions. Since the greatest utility of the VPS should be at larger, impersonal settings where supervisors have less time for systematic individual observation this outcome is encouraging.

Construct validity was assessed by the method of contrasted groups (Cronbach and Meehl, 1954). The disparity between scores of institutionalized and community-based examinees provided one index for construct validity. Experiences related to VPS tasks are clearly more readily available in community workshop/activity center settings than in typical institution prevocational training programs. The VPS appears quite sensitive to these experience-related effects; the mean total VPS scores for 81 community residents was 71% correct, while that for 68 institution residents was 45% (standard deviations were similar, 13.2 and 14.0, respectively). Additionally, as can be seen in Table 1, distributions of scores across the possible range of 0 to 60 were distinctly different for the two groups in the predicted direction, suggesting that extreme scores of a few do not account for the mean differences.

Insert Table 1

The construct validity is being further investigated by a second indicator. Such an indicator is the relating of factor structures between different populations (Kaiser, Hunka, & Bianchini, 1971). The factor structure of the VPS based on the performance of 1) adults from community sites, 2) adults from institution sites, and 3) adolescents from institution sites receiving "special" education will be related.

The factor structure of the VPS was already explored with a sample comprised of <u>both</u> community-based and institutionalized adults and three

salient, orthogonal factors were identified. Factor I, explaining 30 percent of the common variance, groups items that involve imitation and visual discrimination. The second factor, accounting for 27 percent of the common variance, appears to deal with receptive language and more complex cognitive tasks. Factor III, explaining 14 percent of the common variance combines items requiring simple instructions and coordinated hand movements.

Criterion-related validity had long been the near insoluble problem in vocational assessment (Cf. Parker & Hansen, 1976). The current study is using two precisely defined vocational tasks, a small-parts assembly task and a sorting task, much like those used in actual vocational settings. Initial pilot studies indicated reasonably high (negative) correlations between VPS scores and training time to criterion. Although the precise definition of criterion performance has the advantage of easy replicability, there are clear limits to generalizability of findings. Nonetheless this seems a logical starting point. An effort was made to create training situations that were specific enough to allow a standardized procedure to be implemented across different sample sites and different trainers while allowing enough flexibility in training to capture the real clinical nature of the process. This aspect of the study was achieved by piloting the training procedures with a small group of subjects while getting feedback from the trainers. When the training procedures were readjusted to the recommendations of the trainers, 59 severely retarded adults working in sheltered workshops were trained on the two vocational tasks. Total VPS scores were correlated with training time and number of errors to criterion on the two tasks. Pearson correlations between VPS scores and the criterion measures as shown in Table 2 were high, accounting for over 49% of the variance in training time.

Insert Table 2

Conclusions and Significance

Work on this project has led us to believe that tests can be developed which integrate conventional norm-referenced technology with principles of applied behavior analysis, domain-referenced assessment and the learning potential model. It also seems that for severely handicapped populations more tests should be developed using only that population as a validation sample; contrasts to "normal" population are often both pejorative and counterproductive with such an impaired group. If the criterion behavior is clearly specified, then meaningful criterion-related validity studies can be conducted. There is also some evidence that the VPS can lead to aptitude-treatment interaction studies using the Goldfied and Kent (1972) definition of "aptitude" as an operationally defined intervening variable. Factor analysis indicates three orthogonal factors, a) Visual Discrimination/Imitation, b) Receptive Language, c) Physical Dexterity, that may have implication for choice of efficient training procedures.

The fact that VPS scores correlated more highly with supervisors ratings in small, well-staff workshops than large, understaffed institutions gives us hope that the test may prove useful in placement decisions, since many adults will be placed out of institutions in the next decade. Most important is the strong relation between VPS scores and training criteria.

Finally in administering the test to over 150 persons, we developed the sense that this was more than a vocational assessment instrument, that the items were tapping the basic components of intelligent functioning--matching colors, simple two-hand coordination, following simple verbal instructions--without the heavy verbal loading associated with conventional IQ tests. There should be ways where the information gathered on this test can be integrated with other empirical investigations of the development and training of organized, intelligent behavior in young children and with other handicapped populations.

References

Bellamy, G.T., Horner, R.H., & Inman, D.P. Vocational habilitation of severely retarded adults: A direct service technology. Baltimore: University Park Press, 1979.

Budoff, M., & Hamilton, J. Optimizing test performance of moderately and severely mentally retarded adolescents and adults. American Journal of Mental Deficiency, 1976, 81, 49-57.

Feuerstein, R. The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments and techniques. In Studies in cognitive modifiability. Jerusalem, 1972.

Gersten, R., Irvin, L., Bellamy, G.T., Close, D., & Heiry, T. J. Supervisor's Checklist. Unpublished manuscript, University of Oregon, 1979.

Gold, M. Research on the vocational habilitation of the retarded: The present, the future. In N. Ellis (Ed.), International review of research in mental retardation (Vol. 6). New York: Academic Press, 1973, pp. 97-148.

Gold, M.W. Train, don't test. Unpublished manuscript, University of Illinois, 1974.

Goldfried, M.R., & Kent, R.N. Traditional versus behavioral personality assessment. A comparison of methodological and theoretical assumptions. Psychological Bulletin, 1972, 77, 409-420.

Kaiser, H.F., Hunka, S., & Bianchini, J.C. Relating factors between studies based upon different individuals. Multivariate Behavioral Research, 1971, __, 409-422.

Parker, R.M., & Hansen, C.E. Aptitude and achievement tests. In B. Bolton (Ed.), Handbook of measurement and evaluation in rehabilitation. Baltimore: University Park Press, 1976.

Illustration 1

Item Generating Grid

	Single	Double	Movement
Verbal	3 items	6 items	1 item
Model	3 items	6 items	1 item
Prime	3 items	6 items	1 item

Dimension 1

1. Verbal Directions
2. Model/Match-to-Sample
3. Physical Prime

Dimension 2

1. Manipulation on Single Object = SINGLE
2. Manipulation on 2 or More Objects = DOUBLE
3. Physical Movement - or Position: No Objects = MOVEMENT

Difficulty Descriptors -- Dichotomous -- (Sampled representatively across instructional format and number of objects dimensions)

0 = None Present
1 = Minimal Difference
2 = Coordinated (2-hand) Movement
3 = If-Then Rule
4 = Conjunctive Stimulus Control (i.e., 2 Stimulus - Components)
5 = Sequence of Movements Required

Table 1

Percents of Community/Institution-Based TPS Examinees in Different Score Intervals

Score Intervals	Community-Based[a]	Institution-Based[b]
0-15	4.9	25.0
16-30	13.6	29.4
31-45	29.6	35.3
46-60	51.9	10.3

[a] N = 81

[b] N = 68

Table 2

Pearson Correlations Between VPS Scores and Criteria on Vocational Tasks

Criteria	Tasks	
	Assembly	Sort
Time	−.62	−.72
Errors	−.69	−.59

N = 59